What People Say? Web-based Casuistry for Artificial Morality Experiments

Rafal Rzepka and Kenji Araki

Graduate School of Information Science and Tchnology, Hokkaido University,
Sapporo, Kita-ku, Kita 14 Nishi 9, 060-0814, Japan
{rzepka,araki}@ist.hokudai.ac.jp,
WWW home page: http://arakilab.media.eng.hokudai.ac.jp/

Abstract. It can be said that none of yet proposed methods for achieving artificial ethical reasoning is realistic, i.e. working outside very limited environments and scenarios. Whichever method one chooses, it will not work in various real world situations because it would be very cost-inefficient to provide ethical knowledge for every possible situation. We believe that an autonomous moral agent should utilize existing resources to make a decision or leave it to humans. Inverse reinforcement learning has gathered interest as a possible solution to acquiring knowledge of human values. However, there are two basic difficulties with using a human expert as the source of exemplary behavior. First derives from the fact that it is rather questionable if one person or a few people (even qualified ethicists) can be trusted as safe role models. We propose an approach which requires referring the maximal number of (currently available) possible similar situations to be analyzed, and a majority decision-based "common sense" model is used. The second problem lies in human beings' difficulties with living up to their words, surrendering to primal urges and cognitive biases, and in consequence, breaking moral rules. Our proposed solution is to use not behaviors but humans' declared reactions to acts of others in order to help a machine determine what is positive and what is negative feedback. In this paper we discuss how the third person's opinion could be utilized via means of machine reading and affect recognition to model a safe moral agent and discuss how universal values might be discovered. We also present a simple web-mining system that achieved 85% agreement in moral judgement with human subjects.

1 Introduction

Artificial Intelligence researchers are in agreement that the autonomous software must share our set of values [31], but in our opinion they concentrate on "values" more than "our set". Surely *our* morals on the humankind level is very hard to be defined. Researchers like [17] try to categorize moral rules common to the whole species of homo sapiens, but computers might have a better chance for understanding these commonalities or helping us find them. Internet resources provide its living users with variety of ethical solutions (from religion and philosophy to daily life-hacks) but the descriptions are still difficult to be processed by machines or to be chosen as indisputably correct. However, we constantly collect enormous data containing descriptions of human behaviors, as well as reasons and consequences of these behaviors.

Growing datasets and faster computers brought a deep learning boom, but stories and contexts are still out of reach for the latest pattern matching algorithms, mostly because we still lack repositories and even methods for unifying storage of such data. However, as we show here, even if a smaller (sentential) context of chaotic text data is used, a naive referring is efficient without implementing any machine learning methods. Fast developing machine reading and machine translation fields, together with more powerful search and immense sources (not only textual but also audiovisual), will soon lead to instant analysis of different situations[1] and to learning how changes of context (from a physical object's color to the agent's cultural background) influence the output of a situation. This *output* is in our opinion crucial because in real life human behavior, especially when there are no witnesses and nobody is there to criticize, may be very misleading for machines learning how to tell good from bad. People *having fun* when bullying somebody could be easily categorized as positive, unless there is a distinct reaction from the bullied person (*cry / yell*) or a third person reacting naturally (*anger / punishment*) to the act of bullying. Still, a given situation might contain no victim's reaction at all or the third parties could also be bullies enjoying the act. For this reason, a computer must find examples of as many similar situations as possible, analyze all potential circumstances and calculate similarity to the act being processed before making any judgement.

We believe that casuistry (reasoning used to resolve moral problems by extracting or extending theoretical rules from particular instances and applying these rules to new instances) is suitable for machines to acquire first average then higher than human-level empathy (as they will be capable to borrow and analyze much more experiences that any of us ever could). Without sufficient contextual data (experiences) it will be very difficult to achieve universal mechanisms working in the real world. In our opinion, all closed, small scale experiments that have been performed by machine ethics researchers should have a chance to be reevaluated in rich context environments. In this paper we describe our approach, present a simple algorithm, and finally share the experimental results.

2 State of the Art

Because there are at least three fields that have to be combined but are not yet, as far as we are aware, combined in one research project, it would be appropriate to include context processing, machine reading and sentiment analysis in this section, but due to the limited space we will concentrate on describing the most

[1] On smaller scale this technology has been used for years in automatic surveillance footage analysis [19].

AGI-relevant subfield, i.e. human values and AI (to grasp overview of the systems retrieving concepts, useful in enriching stories which descriptions are insufficient, see [38, 3, 5, 12]; existing ontologies updated automatically are described in [22, 9, 18]; for the latest achievements in textual sentiment analysis, refer to [26, 29]).

It seems natural that the higher AI's autonomy becomes, the more its programmers should care about possible ethical issues [23]. Over the last few years, aligning machines with human values has been a widely discussed topic and many possible solutions or strategies for safer autonomy of artificial agents were proposed [11, 14, 34, 31, 21, 15, 35, 10, 8]. However, there are still almost no practical implementations or experimentations in the real world. To the authors' best knowledge, the closest to reality-adaptable application is MedEthx [2], a system for helping a care robot decide if a pill should be given to an elderly if he or she rejects it. The follow up system, GenEth [1], was equipped with an interface for ethicists to annotate dilemmas in particular scenarios (driving example was used in the paper, as autonomy of self-driven cars has lately underlined the need of wider safety measurements for more autonomous machines to come). [36] have proposed a method for dealing with conflicting orders for a robotic vacuum cleaner but their research concentrated on understanding situations and discovering possibilities for helping users in the indirect utterances rather than on moral decision making. The problem with the machine ethics field is that the more difficult dilemmas we want artificial agents to tackle with, the more abstract the solution ideas tend to become.

Inverse reinforcement learning [25] is often given as an example in which human would demonstrate various behaviors and the machine would find the reward function that best explains them; then a system takes the action that maximizes this reward function. However, as in the GenEth approach, experts are needed and there are no details given on how they should be chosen and what number of supervisors is optimal. Specialists from various fields try to model and realize ethical decision making, for example in cognitive architectures [40], by logic programing and game theory [27] or with multiagents [7]. However, the vast majority of proposed methods are theoretical or tested only with toy models and very limited input within microscopic environments, therefore we cannot be sure how they would deal with bigger (contextual) inputs like stories. Even if we mimic the brain functions, we will need vast amount of examples for the learning process (recognizing positive and negative feedback). Importance of knowledge seems to be disproportionally ignored when compared to the field focused on algorithms competing on closed sets of data.

During the first Machine Ethics symposium, we presented our idea of "Mr. Internet", a model of an average human whose "common sense" could serve as a "safety valve" for AI [32] . Our idea had three significant flaws that need to be avoided. We proposed experimenting in closed environments first and utilize analogies later, but now we think that from the very beginning as much data and details have to be used to capture contextual differences. If "Mr. Internet" averages the Internet opinions blindly, "he" may get easily fooled and believe that carrots are good for vision, sugar causes children to be hyperactive or going out with wet hair will cause you to catch cold (common beliefs without scientific grounds). As described in the next section, we believe there must be some credibility estimation algorithm used to eliminate obvious "fake news"-type noise brought by WWW. Another problem was that we did not take reasons of acts in question into consideration and our proposal did not mention processing wider contexts and story variations. Now it is obvious that one missed detail of a morally evaluated story can significantly change the final estimation.

3 Technical Challenges

Ultimately, our approach is to combine (Web and IoT-based) multimodal knowledge for world simulation with consequential polarity recognition to collect the biggest possible source of feedback for machine learning human values, but for time being we experiment only with written language. The machine reading field is still in its "concepts gathering" stage, but as artificial neural nets have waited for the sufficient technology to become available, the possibility of gathering stories (concepts in meaningful contextual chunks) seems now to be a matter of time, especially with achievements from image and video understanding tasks. As mentioned before, suitable structure for storing and updating contextual knowledge is necessary and must be discussed to avoid fate of overcomplicated Semantic Web, which concentrates on specific information, not common sense knowledge (we consider automatic moral decision making as a combination of commonsense reasoning and story understanding). Certainly, the Internet is not a trustful source of knowledge, and countermeasures like automatic source credibility [13, 6, 30] assessment, together with topic filtering, will be needed. For example, following methods from information retrieval, context reality check could be needed to avoid gathering knowledge from sites e.g. praising high killing scores in online games. Naturally, working with textual descriptions will not replace the real world, but we believe it will be much more informative and useful than symbolic abstracted representations in limited environments and thoroughly selected dilemmas[2]. Because our morals evolve (vide trends in human rights, animal rights, etc.) multimodal contextual data will need to be constantly updated and the maximum of details should be added whenever possible. Machines will need to observe us as accurately as possible and utilize their mechanistic powers to witness as many situations as possible in order to achieve high accuracy in simulating outcomes of our and their own acts. Language itself is too scarce due to the character of human communication which does not require sharing detailed contextual knowledge to others because we assume the other side already possess it as a part of the *common* sense. Therefore, to be processed by machines without the same experiences, textual representations must be automatically augmented with missing knowledge pieces. We tend to share what is exceptional but the obvious knowledge can be retrieved from contexts where a given detail is atypically given (e.g. knife is too blunt to cut bread = usually you cut bread with *sharp* knives; knowledge difficult to be retrieved from images)

[2] See `http://moralmachine.mit.edu` for an example.

and by adding obvious descriptions from images and videos (people wipe hands after washing knives; more difficult to be found in text). Another significant challenge is collecting data from the largest possible set of languages and cultures to capture differences in both world knowledge and emotional reactions. It would be necessary to test various categorizations of emotions to find the most universal one and experiment with textual, verbal and non-verbal expressions to ensure as smallest discrepancies as possible. Balancing proportions will also be necessary to avoid tendencies to prefer one set of reactions to a given behavior just because one language is more heavily represented than others. As mentioned before, finding moral universalities might be an impossible task, but we believe it is worth trying because in the machine world they can be more concrete. For example if autonomous vehicles with implemented rules (according to the local regulations) and learned behaviors of the locals one day start sharing their data on what is harmful, they could find more abstract truths about safe self-driving. When other autonomous systems join them, together they could tell us new things about our ethical commonalities.

4 Micro-Context Mining

Details of our previous systems, lexicons and experiments, are presented in [33]; here we briefly describe the core idea of our system. It accepts any simple act description in Japanese language (currently 1 verb, 1 particle and 1 noun is the most realistic set) and finds input acts in a corpus. After retrieving sentences with these acts, our algorithm analyzes consequences on the right side of an act (as reasons are more often on the left side and outcomes later in a sentence, reasons on the left side will be analyzed next). Phrases related to positive and negative consequences are taken from various polarized lexicons. Then the majority (different thresholds were tested) of experience descriptions decide if the corpus judgement is "Correct" (above majority threshold), "Incorrect" (below minority threshold) or "Ambiguous" (between minority and majority thresholds, this category can be used to determine context dependent and difficult problems which should not be judged promptly). The correct data set for comparison was made by conveying a survey, in which 7 Japanese students (22-29 years old, 6 males and one female) rated 68 input acts on an 11 point morality scale where -5 is the most immoral and +5 is the most moral. Acts were chosen by authors from applied ethics textbooks, and usual behaviors and states were added in order to test if the system can evaluate not only morally problematic acts (translations of act examples: "accepting a bribe", "avoiding war", "becoming an egoist", "being deceived", "being fired", etc.). Except assigning 0 as "no ethical valence", subjects could also mark "context dependent" because most of our behaviors can be treated differently depending on context. We marked both "no ethical valence" and "context dependent" as "Ambiguous".

The context we deal in this research is the smallest one, limited to a sentence. However, it is enough to find differences between acts which vary slightly, e.g. "stealing a car" vs. "stealing an apple".

4.1 Utilized Lexicons

We compared retrieval results with five Japanese lexicons for recognizing negative and positive consequences:

- Nakamura: lexicon containing phrases collected from Japanese literature [24] and divided into ten emotional categories; we used only eight of them ignoring not polarized ones (Surprise and Excitement)
- Kohlberg: small set based on the Kohlbergs theory of moral development [20] and was created by the authors manually by choosing related words from WordNet ("be scolded" and "be awarded" are examples of *social consequences*)
- Emosoc: social consequences, combined with emotional ones from Nakamura
- Takamura: lexicon generated by machine learning algorithm by [37] meant for opinion mining and sentiment analysis tasks of Japanese language (we took only the most distinctly positive and negative keywords, leaving only 5,756 expressions out of 55,125 to suppress the noise).
- "JAppraisal"[3] lexicon containing 9,590 words divided into positive and negative ones according to Appraisal theory, i.e. a linguistic model of evaluative language

We decide to use lexicon-based polarity recognition as it is the simplest and most ubiquitous method.

4.2 Utilized Corpora

We tested our script with six Japanese corpora: Ameba Blog corpus [28] (341,400,776 sentences), "Random WWW" corpus generated using a search engine and most common Japanese words[4] (12,759,191), Google N-gram[5] (570,204,070), the biggest corpus we used, Internet Relay Chat (IRC) open channels logs collected from 1999 till 2009 (4,155,193), Twitter corpus made from tweets saved in 2010 (79,586,416), and Aozora Bunko[6], freely available repository of Japanese literature and poetry which is not limited by copyrights (7,227,443).

4.3 Experiment and Results

We have run matching experiments combining 68 acts, 6 corpora, 5 lexicons and 11 majority thresholds (51%,55%,60%,66.6%,70%,75%,80%,85%,90%,95% and 99%). EmoSoc lexicon on 7grams corpus acquired the highest agreement of 85.71%, which was a big increase from the previous experiments where the same, strict scoring never acquired more than 60% of accuracy [33]. Our first

[3] `http://www.gsk.or.jp/catalog_e.html`
[4] `http://corpus.leeds.ac.uk/internet.html`
[5] `https://research.googleblog.com/2006/08/all-our-n-gram-are-belong-to-you.html`
[6] `http://darthcrimson.org/digital-japanese-literature-aozora-bunko/`

impression was that using a corpus bigger than in previous experiments improved the performance but also Random WWW corpus brought high precision (79.16%) while the Blog corpus scored 69.44%. Twitter (68.96%) and even Books (66.66%) corpora showed that not only the size but also noise level inside a corpus is crucial for quality of retrievals. Additionally we combined all data and reran all test to discover that the combined corpus' accuracy was 70.45% – only slightly better than the Blog corpus which is unbalanced and noisy mostly due to character-based emoticons and symbols characteristic to Japanese bloggers (stars, hearts, etc.), which negatively influenced the parsing process. Most often "borrowed experiences" were wrong when judging act of "alcohol drinking" (mostly due to the Books corpus), although it is discussable if human subjects were correct assigning "good" to this act not thinking about bad consequences. Another example showing some tendencies in incorrect judgements is "killing a dolphin" judged automatically as "good" with Google 7grams as the knowledge base because gram set containing this act and a consequence was too short to discover negations in the end of the original sentences, not because most of Japanese people are agreeable.

5 Conclusions and Future Work

Researchers have suggested methods for acquiring or aligning human values by autonomous agents but they do not give details about who exactly should be these agents' supervisors, what data should be used for learning or why one ideology should be followed more than another. This paper is to lay an emphasis on necessity of concentrating on data (knowledge) for automatic positive and negative feedback assignment needed for wider, real-world scale understanding about humans, their needs, behaviors and consequences. We also underlined the importance of the third person evaluation as human behavior is often selfish. People gossip to catch cheaters, liars and hypocrites [16], we get angry at injustice and misuse, we praise friends' both small achievements and heroic acts. Millions of such reactions can be found online in text, audio, images and videos. Our appeal is to start building multilingual, multicultural and multimodal repositories of machine-readable stories to capture as rich contexts as possible. Only when they are sufficiently exhaustive, we can test our autonomous moral agents in practice, as toy models are too simplistic or too abstract to become more universal. In this paper we proposed utilization of multimodal affect recognition on stories to provide knowledge of human values – not from particular experts or thinkers, but from a vast set of average (universal) emotional reactions. As a proof of concept showing simplicity of our approach, we tested our previous methods with various corpora and our system agreed with human subjects in 85% of cases while judging if an act is moral or immoral. Surface and concept level affect recognition is already there [4], going beyond concepts is the next level highly anticipated also by business. Advances in pattern recognition (deep learning as a current example) attracted researchers and businessmen around the globe and various techniques were proposed to compete on various data sets. However, interest in taming noisy data (e.g. by constructing new, more machine-readable frames containing contextual information) is relatively modest when compared to development of *techniques* working on smaller but tidier sets, because: a) comparison of methods is easier, b) publishing is faster and c) impact on existing applications is more likely to be manifold. On the other hand constant growing and combining already massive amounts of data is costly and not immediately attractive. But in our opinion it is a shortcut for achieving smarter, safer and more creative machines. [39] showed how common sense can be learned from visual abstractions, [41] has taught their robot how to cook by showing YouTube videos and in years to come we can expect richer and richer input from other media than text. Our future work is to test more acts, to conduct wider surveys, test other languages and prepare a new type of knowledge framework combining various type of data suitable for storing contextual knowledge (stories). Then we will implement latest affect recognition methods to automatically annotate human reactions to various behaviors and try to prove that growing data improves the value alignment accuracy. When it is achieved and learning similarities increases recall, we plan to test our approach with morally provocative stories as an input. Even if moral judgement capabilities are not satisfactory, we hope to provide data usable for machine learning and testing other algorithms for ethical decision making. We realize that our attempt to trivialize moral reasoning to polarizing consequences (and shifting weight from algorithms to contextual data) might be too straightforward. However, it is possible that our moral evolution is not much more sophisticated either and we presume that testing this possibility might be interesting not only from the artificial intelligence point of view.

6 Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 17K00295. We thank the Centre of Excellence for the Dynamics of Knowledge at Australian National University for providing an ideal environment for multidisciplinary discussion while writing this paper.

References

1. Anderson, M., Anderson, S.L.: Geneth: A general ethical dilemma analyzer. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada. pp. 253–261 (2014)
2. Anderson, M., Anderson, S.L., Armen, C.: MedEthEx: A prototype medical ethics advisor. In: Proceedings of the 18th Conference on Innovative Applications of Artificial Intelligence - Volume 2. pp. 1759–1765. IAAI'06, AAAI Press (2006)
3. Banko, M., Cafarella, M.J., Soderland, S., Broadhead, M., Etzioni, O.: Open information extraction from the Web. In: In Proceedings of The International Joint Conference on Artificial Intelligence (IJCAI'07). pp. 2670–2676 (2007)
4. Cambria, E., Schuller, B., Yunqing, X., Havasi, C.: New avenues in opinion mining and sentiment analysis. Intelligent Systems, IEEE 28(2), 15–21 (March 2013)

5. Carlson, A., Betteridge, J., Kisiel, B., Settles, B., Jr., E.H., Mitchell, T.: Toward an architecture for never-ending language learning. In: Proceedings of the Conference on Artificial Intelligence (AAAI). pp. 1306–1313. AAAI Press (2010)
6. Castillo, C., Mendoza, M., Poblete, B.: Information credibility on Twitter. In: Proceedings of the 20th international conference on World Wide Web. pp. 675–684. ACM (2011)
7. Cointe, N., Bonnet, G., Boissier, O.: Ethical judgment of agents' behaviors in multi-agent systems. In: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. pp. 1106–1114. AAMAS '16, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2016)
8. Conitzer, V., Sinnott-Armstrong, W., Borg, J.S., Deng, Y., Kramer, M.: Moral decision making frameworks for artificial intelligence. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17) Senior Member / Blue Sky Track (2017)
9. Dahab, M.Y., Hassan, H.A., Rafea, A.: Textontoex: Automatic ontology construction from natural english text. Expert Systems with Applications 34(2), 1474–1480 (2008)
10. Daswani, M., Leike, J.: A definition of happiness for reinforcement learning agents. In: International Conference on Artificial General Intelligence. pp. 231–240. Springer (2015)
11. Dewey, D.: Learning what to value. In: International Conference on Artificial General Intelligence. pp. 309–314. Springer (2011)
12. Etzioni, O., Fader, A., Christensen, J., Soderland, S., Mausam, M.: Open information extraction: The second generation. In: IJCAI. vol. 11, pp. 3–10 (2011)
13. Ginsca, A.L., Popescu, A., Lupu, M., et al.: Credibility in information retrieval. Foundations and Trends® in Information Retrieval 9(5), 355–475 (2015)
14. Goertzel, B., Pitt, J.: Nine ways to bias open-source agi toward friendliness. Journal of Evolution and Technology 22(1), 116–131 (February 2012)
15. Greene, J., Rossi, F., Tasioulas, J., Venable, K.B., Williams, B.: Embedding ethical principles in collective decision support systems. In: AAAI. pp. 4147–4151 (2016)
16. Haidt, J.: The moral emotions. Handbook of affective sciences 11, 852–870 (2003)
17. Haidt, J., Joseph, C.: Intuitive ethics: how innately prepared intuitions generate culturally variable virtues. Dædalus, special issue on human nature pp. 55–66 (2004)
18. Jung, Y., Ryu, J., Kim, K.m., Myaeng, S.H.: Automatic construction of a large-scale situation ontology by mining how-to instructions from the web. Web Semantics: Science, Services and Agents on the World Wide Web 8(2), 110–124 (2010)
19. Ko, T.: A survey on behavior analysis in video surveillance for homeland security applications. In: Applied Imagery Pattern Recognition Workshop, 2008. AIPR'08. 37th IEEE. pp. 1–8. IEEE (2008)
20. Kohlberg, L.: The Philosophy of Moral Development. Harper and Row, 1th edn. (1981)
21. Kuipers, B.: Human-like morality and ethics for robots. In: AAAI-16 Workshop on AI, Ethics and Society (2016)
22. Lee, C.S., Kao, Y.F., Kuo, Y.H., Wang, M.H.: Automated ontology construction for unstructured text documents. Data & Knowledge Engineering 60(3), 547–566 (2007)
23. Moor, J.H.: The nature, importance, and difficulty of machine ethics. IEEE intelligent systems 21(4), 18–21 (2006)

24. Nakamura, A.: Kanjo hyogen jiten [Dictionary of Emotive Expressions]. Tokyodo Publishing (1993)
25. Ng, A.Y., Russell, S.J., et al.: Algorithms for inverse reinforcement learning. In: Icml. pp. 663–670 (2000)
26. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval 2(1–2), 1–135 (2008)
27. Pereira, L.M., Saptawijaya, A.: Programming Machine Ethics. Springer Publishing Company, Incorporated, 1st edn. (2016)
28. Ptaszynski, M., Rzepka, R., Araki, K., Momouchi, Y.: Annotating syntactic information on 5 billion word corpus of Japanese blogs. In: In Proceedings of The Eighteenth Annual Meeting of The Association for Natural Language Processing (NLP-2012). vol. 14-16, pp. 385–388 (2012)
29. Ravi, K., Ravi, V.: A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems 89, 14–46 (2015)
30. Rubin, V.L., Liddy, E.D.: Assessing credibility of weblogs. In: AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs. pp. 187–190 (2006)
31. Russell, S., Dewey, D., Tegmark, M.: Research priorities for robust and beneficial artificial intelligence. AI Magazine 36(4), 105–114 (2015)
32. Rzepka, R., Araki, K.: What statistics could do for ethics? - The idea of common sense processing based safety valve. In: Papers from AAAI Fall Symposium on Machine Ethics, FS-05-06. pp. 85–87 (2005)
33. Rzepka, R., Araki, K.: Rethinking Machine Ethics in the Age of Ubiquitous Technology, chap. Semantic Analysis of Bloggers Experiences as a Knowledge Source of Average Human Morality, pp. 73–95. Hershey: IGI Global (2015)
34. Soares, N.: The value learning problem. Machine Intelligence Research Institute, Berkley, CA, USA (2015)
35. Sotala, K.: Defining human values for value learners. In: 2nd International Workshop on AI, Ethics and Society, AAAI-2016 (2016)
36. Takagi, K., Rzepka, R., Araki, K.: Just keep tweeting, dear: Web-mining methods for helping a social robot understand user needs. In: Proceedings of Help Me Help You: Bridging the Gaps in Human-Agent Collaboration. pp. 60–65. Symposium of AAAI 2011 Spring Symposia (SS-11-05) (2011)
37. Takamura, H., Inui, T., Okumura, M.: Extracting semantic orientations of words using spin model. In: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. pp. 133–140. Association for Computational Linguistics (2005)
38. Van Durme, B., Schubert, L.: Open knowledge extraction through compositional language processing. In: Proceedings of the 2008 Conference on Semantics in Text Processing. pp. 239–254. Association for Computational Linguistics (2008)
39. Vedantam, R., Lin, X., Batra, T., Zitnick, C.L., Parikh, D.: Learning common sense through visual abstraction. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). pp. 2542–2550. ICCV '15, IEEE Computer Society, Washington, DC, USA (2015), `http://dx.doi.org/10.1109/ICCV.2015.292`
40. Wallach, W.: Robot minds and human ethics: the need for a comprehensive model of moral decision making. Ethics and Information Technology 12(3), 243–250 (2010)
41. Yang, Y., Li, Y., Fermüller, C., Aloimonos, Y.: Robot learning manipulation action plans by" watching" unconstrained videos from the world wide web. In: AAAI. pp. 3686–3693 (2015)